Floor Discussion

Michael Ikeda
Bureau of the Census

The first paper, presented by Danny Pfefferman (Hebrew University of Jerusalem), proposed a method for estimating trends for small areas. One difficulty in small area trend estimation is that survey errors are often strongly autocorrelated and this produces spurious trends under the standard X-11 model. The proposed method deals with this problem by modelling both the sampling error autocorrelations and the population values.

The second paper, presented by Gauri Datta (University of Georgia), outlined a hierarchical Bayes time series model for small area (state-level) labor force estimates. The proposed model combines cross-sectional and time series information. A comparison of the proposed model with the current model for state-level labor force estimates suggested that the proposed model provides better estimates.

The discussant, Richard Tiller (Bureau of Labor Statistics), began with an overview of the Pfefferman paper. He noted three key implications of the problem of autocorrelated sampling errors: design-based estimates of sampling error autocorrelations are important, model-based signal extraction is a useful analysis tool, modifications to X-11 need to be considered. Two possibilities for X-11 are (I) to modify the X-11 filter, or (ii) to apply a signal extraction filter to remove sampling error from the data before feeding it to X-11 for trend estimation. Tiller noted that more research needs to be done on trend filters.

Richard Tiller continued with an overview of the Datta paper. He noted that key issues for the proposed model include: which areas to combine, model diagnostics, operational considerations, and quantifying uncertainty.

Alan Zaslavsky (Harvard University) opened the floor discussion by asking whether the trend filtering approach can be generalized across different series. He noted that for high-frequency noise you don't really care about over filtering, but for low-frequency noise the signal to noise ratio depends on the estimate of the trend, and explicit modeling may be more necessary. Richard Tiller responded that there are different ways of estimating the trend. BLS tries to first estimate the true series and then applies X-11 to this to estimate the trend. Danny Pfefferman responded that the same problem occurs for the X-11 method. He noted that even if the standard error model is misspecified it doesn't have a great effect on the estimates produced by his proposed method.

Danny Pfefferman remarked that, in his experience, accounting for past data was more important than pooling cross-sectional data. He asked whether Datta had looked at the relative contributions of the time series and cross-sectional information. Gauri Datta replied that this hadn't been done yet. It will be looked at in the future. Bill Bell (Bureau of the Census) commented that, in a 1993 ARC paper by Ghosh and Nangia (that used a model similar to the one used by Datta, Lahiri, and Lu), most of the improvement came from the cross-sectional information.

Bill Bell commented on Pfeffermann's paper that it is important to consider the range of possible

trends consistent with any given model for the observed data. A wide range of trend estimates from quite rough to fairly smooth is generally possible depending on what is assumed about the components (as is illustrated in a 1978 *Biometrika* paper by Tiao and Hillmer).

Partha Lahiri (University of Nebraska) asked about the estimates of error under the model used by Pfefferman. He noted that if the MLEs are simply plugged back into the model for the purpose of estimating the error variance of the trend estimates, that this does not account for the variance due to the estimation of the model parameters. Danny Pfefferman responded that this additional source of error was ignored, and that a Bayesian approach could be used to account for it. However, this extra variance is of a lower order than the other variance components. Bill Bell noted that one other advantage to using a Bayesian approach is that it would get around the problem Pfeffermann encountered of getting zero for the MLE of a variance.